ZACKARY SCALYER
zack0179@gmail.com
www.linkedin.com/in/ZackaryScalyer
github.com/zack0179/ZackaryScalyer

DATA SCIENTIST | MACHINE LEARNING

Data Scientist and Marine Corps Veteran with knowledge and experience in statistical and data mining techniques: Linear/Generalized Linear Regression, Categorical Analysis, Bayesian Analysis, blocking, random effects modeling, dimension/category reduction, text/web scraping, down/up sampling, imputation, and signal processing. Experience querying databases and using computer languages: R, Python, MATLAB, SAS, and SLQ to manipulate data and draw insights from large data sets. Strong knowledge creating and using advanced machine learning algorithms and statistics: TensorFlow, Pytorch, Keras, Decision Tree, Random Forest, Neural Network, Gaussian Mixture Model/K-means Clustering, Multicategory Logit Regression, Monte Carlo simulation, bagging analysis, and experimental design. Proficient in assessing the quality and accuracy of new data sources and data gathering techniques. Skilled in object-oriented programing automation for applying and assessing data models and algorithms. Excellent communication skills that permit collaboration with company stakeholders throughout the organization to identify opportunities for leveraging company data to drive business solutions. Demonstrated organization and time management ability during numerous projects in diverse environments.

- Algorithm Development
- Object-Oriented Programming
- Data Management
- Advanced Statistical Modeling
- Automation
- Workflow Documentation
- Hyperparameter Tuning
- Data Analysis
- Collaboration

EDUCATION

Master of Science, Applied Statistics | Villanova University, Villanova, PA
Bachelor of Science, Kinesiology & Exercise Science | The Pennsylvania State University, Berks Campus, Reading, PA

RELEVANT PROJECTS

Undergraduate Thesis

- Developed and presented a research problem, paying attention to scientific and philosophical justifications
- Formulated a logical, step-by-step design for data collection and informed consent form
- Prepared a completed research thesis of 4 chapters, ensuring all scheduled deadlines were met
- Collected and analyzed Electromyography (EMG), Ultrasound, and muscle force subject data using MATLAB

Regression Methods

- Conducted regression model building, analysis, and considered all assumptions
- Addressed non constant variance and normality with Weighted Least Squares and Robust Regression
- Collected and merged open source data using Python and Google's BigQuery API (SQL)
- Created R functions for automated model building and comparison through nested-data structure

Data Mining

- Collaborated with team of 5 peers on a classification task in a business setting
- Integrated R codes and business rules into comprehensive annotated code
- Produced data objects at each data step and independent scoring function for unseen data
- Considered the quality of propriety survey data, applied Gaussian Mixture Model clustering
- Fitted, tuned, and compared Artificial Neural Network, Decision Tree, and Random Forest
- Directed low target prevalence with receiver operating curve (ROC) optimum thresholds and down sampling

Categorical Data Analysis

- Managed multicategory logit model building with forward stepwise procedure in R
- Generated R functions to output model effects and hypothesis tests comparable to SAS
- Defined improved class prediction from predicted probabilities, compared to undocumented methods in R
- Examined model sensitivity and precision measure in multiclass context

Bayesian Statistics

- Estimated National Hockey League team power and home ice parameters with Metropolis Hastings algorithm
- Implemented Empirical Bayes methods for choosing prior distributions and hyperparameters
- Automated posterior time-series plots to optimize proposal domains with interactive output in R Markdown
- Calculated credible intervals for team's power parameter estimation and rank

PROFESSIONAL EXPERIENCE

The Pennsylvania State University, Berks Campus, Reading, PA
Biomechanics Research Studies **2016 – 2018**

- Received a research grant to conduct my undergraduate thesis
- Investigated in vivo muscle function by integrating and analyzing Ultrasound imaging, Electromyography (EMG), and load cell data using MATLAB
- Presented research findings as a poster session to an international audience
- Developed instructional documentation for Vicon Nexus II Motion Analysis System
- Aided students in data collection and faculty in class demonstrations

The Pennsylvania State University, Berks Campus, Reading, PA
Physics Department Research Studies **2017 – 2018**

- Conducted collaborative research on remote repository using Git
- Implemented Monte Carlo simulation and data visualization with Python on Linux Redhat server
- Developed instructional documentation for remote server and GitHub access
- Research findings were presented during the campus Science Division Poster Session

The Pennsylvania State University, Berks Campus, Reading, PA **2017**
Engineering Department Research Assistant

- Designed and administered subject data collection protocol in validation experiment
- Created MATLAB scrip to process and store subject data in a struct, maintaining data accountability and participant confidentiality
- Implemented variances tests, ANOVA, and intraclass correlation coefficient in MATLAB in a systematic way using a data struct

The Pennsylvania State University, Berks Campus, Reading, PA **2016 – 2017**
Learning Center Tutor & Assistant Supervisor

- Presented group classes to oncoming engineering students for annual four-week math preparation program in R programing and regression sponsored by National Science Foundation (NSF) Summer Engineering Ahead
- Conducted individual and group tutoring 5-15 hours per week in the learning center for Math, Science, and Statistics
- Advised Psychology and Kinesiology majors in thesis research/planning survey methods in SPSS
- Led STAT 200 hypotheses testing lecture and STAT 100 introduction to probability lectures

The Pennsylvania State University, Berks Campus, Reading, PA **2016**
Chemistry Department Research Assistant

- Identified errors in large data, cleaned, reduced, transformed, and extracting meaningful data using R
- Implemented Linear and Generalized Linear Regression for predictive modeling of classroom performance
- Researched findings were presented in a poster and podium session
- Provided statistical programing consulting for fallowing assistants beyond employment tenure

United States Marine Corps |Various Locations **2008 – 2013**
Infantry Rifleman

- Served as Assistant Patrol Leader (APL) for numerous training evolutions in six different countries
- Martial Art Instructor conducting over 158 hours of formal instruction, testing 39 students
- Received Certificate of Commendation while serving as Ground Safety NCO at Jungle Warfare Training Center

PRESENTED RESEARCH

PSU Berks Science Division Poster Session 2018|The Structure of the Building Blocks of the Universe
41st American Society of Biomechanics Annual Conference 2017|Optimal Fascicle Length Changes Based on Submaximal Force or Activation
18th Annual Undergraduate Research & Creativity Conference 2017|Organic Chemistry Predicting Outcomes
Biennial Conference on Chemistry Education (BCCE) 2016|Linear and Generalized Linear Modeling of Organic Chemistry

ACADEMIC AWARDS

Villanova's Applied Statistics Comprehensive Exam Pass with Distinction Honor|2020
The Penn State Berks Excellence in Statistics Award|2018
Franco Undergraduate Research Award|2017
Science Peer Tutor Award: Kinesiology |2017
Outstanding Science Internship/Co-op Student Award|2017
Science Peer Tutor Award: Mathematics|2016